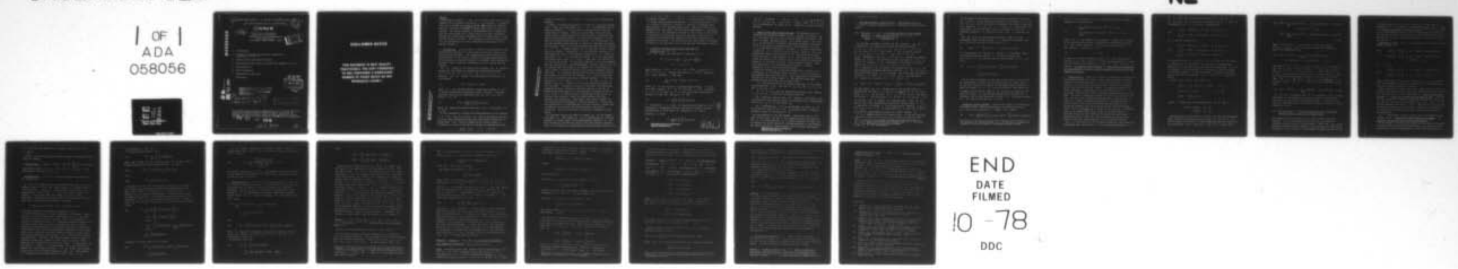
AD-A058 056 BROWN UNIV PROVIDENCE R I DIV OF APPLIED MATHEMATICS F/G 12/1
APPROXIMATION METHODS FOR THE MINIMUM AVERAGE COST PER UNIT TIM--ETC(U)
MAY 78 H J KUSHNER N00014-76-C-0279

UNCLASSIFIED NL

1 OF 1
ADA
058056



END
DATE
FILMED
10 -78
DDC

(6)

APPROXIMATION METHODS FOR THE MINIMUM AVERAGE COST PER UNIT TIME PROBLEM WITH A DIFFUSION MODEL

(12)

(11) 12 May 78

(12) 21p.

(10) Harold J. Kushner*
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

(15) N00014-76-C-0279,
AFOSR-76-3063

*This research was partially supported by the Air Force Office of Scientific Research under AF-AFOSR 76-3063, in part by the National Science Foundation under NSF-Eng73-03846-A-03 and in part by the Office of Naval Research under N00014-76-C-0279-P0002.

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DDC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Abstract

Approximation methods for the minimum average cost per unit time problem with a controlled diffusion model is treated. In order to work with a bounded state space, ~~we use the~~ a reflecting diffusion model of Strook and Varadhan is used, although other models can also be treated. The control problem is approximated by an average cost per unit time problem for a Markov chain, and weak convergence methods are used to show convergence of the minimum costs to that for the optimal diffusion. The procedure is quite natural and allows the approximation of many interesting functionals of the optimal process.

1. Introduction. In this paper, we develop an approximation and computational approach to a particularly difficult class of stocastic control problems. The computational problem leads to the approximation of the original process and optimization problem by an interesting and simpler sequence of processes and optimization problems, which yields much information on the original optimal process.

Let $w(\cdot)$ denote an R^r-valued Wiener process, let $\mathscr{U}$ denote a compact set and define the bounded and continuous functions $f(\cdot,\cdot)\colon R^r \times \mathscr{U} \to R^r$; $k(\cdot,\cdot)\colon R^r \times \mathscr{U} \to R$; $\sigma(\cdot)\colon R^r \to r \times r$ matrices. Let $x(\cdot)$ denote a non-anticipative solution to the Itô equation

$$dx = f(x,u)dt + \sigma(x)dw, \tag{1}$$

where $u(\cdot)$ is a non-anticipative (always with respect to $w(\cdot)$) $\mathscr{U}$-valued progressively measurable control function. For typographical simplicity we sometimes write x_s for $x(s)$, etc.. Define $\gamma^u(\cdot)$ by

$$\gamma^u(x) = \overline{\lim_{t\to\infty}}\, E_x^u \frac{1}{t}\int_0^t k(x_s,u_s)ds, \tag{2}$$

where E_x^u denotes the expectation when $x_0 = x$ and control $u(\cdot)$ is used.

We are interested in finding good approximations to the infimum $\overline{\gamma}$ of $\gamma^u(x)$ over all controls $u(\cdot)$, and to the optimal control, and also other information concerning the optimal trajectory, in cases where $\gamma^u(x)$ does not depend on the initial state x. Furthermore, we want to be able to compute the approximation and obtain the additional information by using practical computational methods.

A number of difficulties stand in the way of a practical computation. First, the state space R^r of $x(\cdot)$ is unbounded and the control problem (1) - (2) will have to be modified so that the state space is bounded. This is a particularly ticklish point, since we want a modification which yields usable information concerning the original problem. In particular situations, a great deal of attention must be devoted to this. For definiteness, we use the bounded process defined in Section 4, although many others are possible. Next, we have not assumed very much about the system (1). If $\gamma^u(\cdot)$ actually depends on x, then very little is known about the problem. Fortunately, for many problems (perhaps the most important ones) we can restrict attention to $u(\cdot)$ which are stationary ($u(\cdot)$ is a stationary process), or to the stationary pure Markov case (where $u_t = u(x_t)$). Even then, the solution to (1) may not be unique. In practical problems, it is often demanded that the system have a certain robustness. Criteria such as (2) are of interest when the system is to operate over a long period of time, usually of uncertain duration and with an uncertain initial condition. It is usually desired that the control be stationary pure Markov, and that for the controls $u(\cdot)$ in the class which are to be considered there be an invariant measure μ^u, and the measures of $x(t)$ tend to μ^u as $t \to \infty$ for each $x = x_0$. In certain cases (e.g., Kushner [1]) one can restrict attention to such controls. In general, little is known about the continuous parameter problem, and many of the difficulties in the way of establishing convergence of a computational procedure are due to this. Also, it is usually hard to approximate problems over an infinite time interval, unless the approximation and limit processes are stationary. Furthermore, the ergodic subsets for each approximation may depend on the approximation. In any case, the procedures to be developed here are very natural, provide much information, and do give the desired results under broad conditions. We will later make an additional assumption on the system.

Our approach follows the ideas in Kushner [2], [3] and Kushner and DiMasi [4]. The problem (1), (2) is approximated by a control problem on a Markov chain (with approximation parameter h), and weak convergence methods are used to show that certain interpolations of the sequence of approximating chains converge weakly to an optimal process. The method yields a great deal of information on the optimal process; e.g., invariant measures and joint distributions.

A formal dynamic programming approach to the optimization of (1), (2) is given in Section 2, Section 3 argues for a "computational approximation" and a bounded state space. The actual form of the bounded state space model, the Strook-Varadhan model of a reflected diffusion [5], is discussed in Section 4. This model is used partly for the sake of specificity and partly because it allows us to illustrate some interesting features of the weak convergence and boundary time scaling. The actual discrete state model is developed in Section 5 and Sections 6 and 7 give the weak convergence results.

2. A Dynamic Programming Sufficient Condition for Optimality for (1), (2).

Let $\mathscr{L}^u$ denote the differential generator of (1):

$$\mathscr{L}^u = \sum_i f_i(x,u) \frac{\partial}{\partial x_i} + \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

$$a(\cdot) = \sigma(\cdot)\sigma(\cdot)'/2.$$

When evaluating $\mathscr{L}^u F(\cdot)$ at t,ω, for a $C^2(R^r)$ function $F(\cdot)$, set $x = x_t$, $u = u_t$. Suppose that there is a $C^2(R^r)$ function $V(\cdot)$ and a constant $\bar{\gamma}$ such that

$$\inf_{\alpha \in \mathscr{U}} [\mathscr{L}^\alpha V(x) + k(x,\alpha) - \bar{\gamma}] = 0, \tag{3}$$

where $\mathscr{L}^\alpha$ is now treated as a parametrized operator. If there is a Borel function $\bar{u}(\cdot)$ on R^r such that $\alpha = \bar{u}(x)$ minimizes at x in (3) for each $x \in R^r$, and to which there corresponds a process (1) such that $E_x^u V(x_t)/t \to 0$, then

$$\bar{\gamma} = \lim_{t\to\infty} \frac{1}{t} E_x^{\bar{u}} \int_0^t k(x_s, \bar{u}_s)\, ds. \tag{4a}$$

If, in addition, $v(\cdot)$ is any $\mathscr{U}$-valued non-anticipative (ω,t) progressively measurable function (henceforth called a control) corresponding to which there is a solution to (1), and if $\frac{1}{t} E_x^v V(x_t) \to 0$, then

$$\bar{\gamma} \le \underline{\lim}_{t\to\infty} \frac{1}{t} E_x^v \int_0^t k(x_s, v_s)\, ds, \tag{4b}$$

and $\bar{u}(\cdot)$ is optimal with respect to such $v(\cdot)$ in the sense that $\gamma^{\bar{u}} \leq \gamma^{v}$ for any x_0 either fixed or random. Under $\bar{u}(\cdot)$ or $v(\cdot)$, (1) is homogeneous, but there is not necessarily a unique invariant measure.

3. Bounded State Space Approximations. The approximation and computational method developed in [2] is roughly as follows. Let $u(\cdot)$ be fixed, and let it be a function only of the state x. We derive a family (parametrized by h) of Markov chains. For fixed $u(\cdot)$, the sequence of (suitable) continuous parameter interpolations of the chains converge weakly to the solution to (1), as $h \to 0$, under broad conditions. For each h, we have a controlled (indexed by $u(\cdot)$) family of Markov chains. Optimize, using the appropriate Markov chain version of (2), and obtain the minimum value function for each chain. As $h \to 0$, the sequence of minimum values converges to the infimum, over a large class of comparison controls, of the value function of the original problem. Also, many properties of the approximations converge to similar properties of the limiting optimal process.

Since our interest is in feasible computations, as well as in convergence, it is necessary that for each h the state space of the approximating chain be finite. This requirement necessitates revision of the original system (1). The following are among several possibilities that can be dealt with.

(i) The state space may be naturally bounded, in that there are bounded sets G_0, G_1 such that if $x_0 \in G_0$, then $x_t \in G_1$ for all t and all $u(\cdot)$.

(ii) If $x_0 \in G_0$, then the approximating Markov chain remains in G_1, for all h, under the optimizing controls.

(iii) Impulsive control terms ([2], Chapter 8) are added to the cost function, such that the state is guaranteed to be "impulsively" driven into G_0, if it ever leaves G_1.

(iv) A bounded set G can be introduced, such that x_t is not allowed to leave $\bar{G} = G + \partial G$. To guarantee this, a suitable boundary process is introduced on ∂G.

For concreteness in the development, a particular form of (iv) will be dealt with. We let G be a hyper-rectangle and suppose that x_t is reflected from ∂G. A hyper-rectangle is chosen only to simplify the specification of the approximation on the boundary. Any region for which a specification with the proper convergence properties exists can be chosen.

4. The Submartingale Problem of Strook and Varadhan [5] in G.

In order to assure ourselves that the reflection is well defined, assume

(A1) for each i, $a_{ii}(x)$ is strictly positive on the boundary planes of $\bar{G}$ which are parallel to $\{x: x_i = 0\}$, where $x_i = i^{th}$ component of x.

We introduce a boundary control and cost function. Let $\mathscr{U}_0$ be a compact set, and define the bounded continuous functions $\gamma(\cdot,\cdot): \partial G \times \mathscr{U}_0 \to R^r$; $k_0(\cdot,\cdot): \partial G \times \mathscr{U}_0 \to R$; $\rho(\cdot): \partial G \to [0,1]$. Let the vector $\gamma(x,\alpha)$ with origin x point strictly interior to G for each $x \in \partial G$ and $\alpha \in \mathscr{U}_0$. For $A \subset R^r$, set $I_A(x)$ = indicator of set $\{x: x \in A\}$, let $x(\cdot)$ denote the generic element of $C^r[0,\infty)$ (R^r-valued continuous functions on $[0,\infty)$) as well as the solution to (1). Hopefully, no confusion will arise. Define $C_G^r = C^r[0,\infty) \cap \{x(\cdot): x_t \in \bar{G}, \text{ all } t < \infty\}$ and $\mathscr{G}_t$ = σ-algebra on C_G^r induced by the projections x_s, $s \le t$. For this reflecting diffusion, admissible controls $u(\cdot)$ are $\mathscr{U}$-valued when the process state $x_t \in \partial G$, and are $\mathscr{U}_0$-valued when the process state++ $x_t \in \partial G$. For+ $q(\cdot,\cdot) \in C^{2,1}(\bar{G} \times [0,\infty))$ and admissible $u(\cdot)$, define the function $F_q^u(\cdot,\cdot)$ on $C_G^r[0,\infty)$ by

$$(5) \quad F_q^u(x(\cdot),t) = q(x_t,t) - q(x_0,0) - \int_0^t [\frac{\partial}{\partial s} + \mathscr{L}^u] q(x_s,s) I_G(x_s) ds.$$

For the moment, let $u(\cdot)$ depend only on the current state x. Suppose that for some $y \in \bar{G}$, there is a measure P_y^u on C_G^r such $P_y^u\{x_0 = y\} = 1$ and for each $q(\cdot,\cdot)$ in $C^{2,1}(\bar{G} \times [0,\infty))$ for which $\rho(x)q_t(x,t) + \gamma'(x,u(x))q_x(x,t) \ge 0$ for all $x \in \partial G$, and all $t \ge 0$, the process $\{F_q^u(\cdot,t), \mathscr{B}_t, P_y^u\}$ is a submartingale. Then P_y^u is said to solve the submartingale problem for initial value y. If, in the above, the vector y can be replaced by a measure ν_0 on $\bar{G}$, and $P_{\nu_0}^u\{x_0 \in \Gamma\} = \nu_0(\Gamma)$ for each Borel set Γ, then $P_{\nu_0}^u$ is said to solve the submartingale problem for initial measure ν_0.

If $u(\cdot)$ depends only on the current state x, then the solution

---

+ $C^{2,1}$ is the set of uniformly bounded continuous functions on $\bar{G} \times [0,\infty)$ whose derivatives up to second order in x and first in t, are continuous and uniformly bounded.

++ and u_t is $\mathscr{G}_t$ measurable.

to the submartingale problem gives the desired reflected diffusion, and $\gamma(x,u(x))$ is the average "direction of reflection" at $x \in \partial G$, and $\rho(x)$ is a scale factor which determines the relative time that $x(\cdot)$ spends on ∂G ([2], [3], [5]). Since $\rho(\cdot)$ only affects the time scale, and not the costs ([3], [2], Chapter 10), for our modelling purpose it is sufficient to set $\rho(x) \equiv 1$, which we will do.

Let P_y^u solve the submartingale problem. There is a non-decreasing scalar valued process $\rho(\cdot)$, which only increases when $x_t \in \partial G$, and is such that for the above $q(\cdot,\cdot)$

$$(6) \qquad F_q^u(x(\cdot),t) - \int_0^t [q_s(x_s,s) + \gamma'(x_s,u_s)q_x(x_s,s)]d\mu_s$$

is a martingale (with respect to $\{P_y^u, \mathscr{B}_t\}$). Furthermore, there is a standard Wiener process[+] $w(\cdot)$ such that under P_y^u, $(x(\cdot),u(\cdot),\mu(\cdot))$ are non-anticipative with respect to $w(\cdot)$ and w.p.1.

$$(7) \qquad x_t = y + \int_0^t f(x_s,u_s)I_G(x_s)ds + \int_0^t \sigma(x_s)I_G(x_s)dw_s$$
$$+ \int_0^t I_{\partial G}(x_s)\gamma(x_s,u_s)d\mu_s.$$

For the control problem, we may wish to deal with a larger class of (admissible) controls than the stationary pure Markov class. We can still speak of a solution to the submartingale problem, but then the measure P_y^u or $P_{\nu_0}^u$ must be defined on the appropriate σ-algebra on the product space of C_G^r and the path space for the control process. If this extended submartingale problem has a solution, then the non-decreasing process $\mu(\cdot)$ and Wiener process $w(\cdot)$ will still exist and (6), (7) hold.

<u>A modified control problem</u>. Suppose that there is a solution to the submartingale problem corresponding to admissible control $u(\cdot)$, and initial condition y. Define $\gamma^u(y)$ now by

$$(8) \qquad \gamma^u(y) = \overline{\lim_{t\to\infty}} \frac{1}{t} E_y^u\{\int_0^t k(x_s,u_s)I_G(x_s)ds + \int_0^t k_0(x_s,u_s)I_{\partial G}(x_s)d\mu_s\}.$$

[+]To construct the Wiener process $w(\cdot)$, we may have to augment the probability space by adding an independent Wiener process.

Since $\rho = 1$, we can set $\mu_s = s$. The formal dynamic programming equation (3) is replaced by

$$\text{(9)} \qquad \begin{aligned} &\inf_{\alpha \in \mathscr{U}} [\mathscr{L}^{\alpha} V(x) + k(x,\alpha) - \bar{\gamma}] = 0, \quad x \in G, \\ &\inf_{\alpha \in \mathscr{U}_0} [V_x'(x)\gamma(x,\alpha) + k_0(x,\alpha) - \bar{\gamma}] = 0, \quad x \in \partial G, \end{aligned}$$

where $V(\cdot)$ is now assumed to be bounded. If there is a solution to the submartingale problem corresponding to admissible control $v(\cdot)$ and initial condition y, and also a smooth function $V(\cdot)$ and constant $\bar{\gamma}$ solving (9), then

$$\text{(10)} \qquad \bar{\gamma} \leq \gamma^{v}(y).$$

If there is a Borel admissible control $\bar{u}(\cdot)$ which attains the infimum in (9), and for which the submartingale problem has a solution for each initial condition x, then $\bar{\gamma} = \bar{\gamma}^{u}(y)$ and $\bar{u}(\cdot)$ is optimal. We emphasize that although (9) will serve as the basis of our approximation, it need not have a solution of any sort for our method to be valued.

5. Discretization. There are a number of techniques for getting an approximating sequence of Markov chain control problems with the correct convergence properties. We use the method in [2] mainly because it is relatively straightforward, fairly well understood and we can refer to existing results. The method is based on a finite difference approximation with difference interval h. A particular (but natural) finite difference approximation to (9) is used. It makes no difference whether or not (9) has a smooth solution, for the finite difference approximation is not used to solve (9). After a suitable rearrangement, the coefficients of certain terms in the finite difference approximation will be transition probabilities for an approximating controlled Markov chain. This is the only use to which (9) will be put. The method gives us an approximating chain simply and automatically. A detailed outline of the method and of some of the convergence properties will be given, but many of the details which can be found in the basic references [2], [3], [4] will be omitted.

Let e_i = unit vector in i^{th} coordinate direction, and assume for convenience that each side of G is an integral multiple of h.

Let G_h denote the finite difference grid on G, and set $\partial G_h =$ $\overline{G}_h - G_h$, where $\overline{G}_h$ is the finite difference grid on $\overline{G}$. Now, let us discretize (9). On ∂G, use the approximation

$$
(11) \quad
\begin{aligned}
V_{x_i}(x) &\to [V(x+e_i h) - V(x)]/h, \quad \text{if} \quad \gamma_i(x,\alpha) \geq 0 \\
V_{x_i}(x) &\to [V(x) - V(x-e_i h)]/h, \quad \text{if} \quad \gamma_i(x,\alpha) < 0.
\end{aligned}
$$

In G, use the approximation

$$
(12) \quad
\begin{aligned}
V_{x_i}(x) &\to [V(x+e_i h) - V(x)]/h, \quad \text{if} \quad f_i(x,\alpha) \geq 0 \\
V_{x_i}(x) &\to [V(x) - V(x-e_i h)]/h, \quad \text{if} \quad f_i(x,\alpha) < 0 \\
V_{x_i x_i}(x) &\to [V(x+e_i h) + V(x-e_i h) - 2V(x)]/h^2.
\end{aligned}
$$

The approximations for $V_{x_i x_j}(x)$, $i \neq j$, are long, and the reader is referred to [2], Chapter 6.2 for one set of possibilities. Simply to avoid writing these down here, we suppose that $\sigma(x)\sigma'(x)$ is diagonal. This assumption is not required by anything except our current laziness. It does not affect the outcome, only the precise form of the functions $Q_h(\cdot,\cdot)$ and $p^h(\cdot,\cdot)$ introduced below.

Define $Q_h(x,\cdot)$, $\Delta t^h(x)$ and $\overline{Q}_h(x)$ by

$$
\begin{aligned}
Q_h(x,\alpha) &= h \sum_i |f_i(x,\alpha)| + \sum_i \sigma_i^2(x), \quad && x \in G_h, \\
Q_h(x,\alpha) &= \sum_i |\gamma_i(x,\alpha)| \quad &&, \quad x \in \partial G_h \\
\overline{Q}_h(x) &= \sup_\alpha Q_h(x,\alpha),
\end{aligned}
$$

(where α ranges over the appropriate set $\mathcal{U}$ or $\mathcal{U}_0$),

$$
\begin{aligned}
\Delta t^h(x) &= h/\overline{Q}_h(x) \quad \text{on} \quad \partial G_h, \\
&= h^2/\overline{Q}_h(x) \quad \text{on} \quad G_h.
\end{aligned}
$$

Approximating the derivatives in (9) by (11)-(12) and rearranging terms yields the following equation, where $\overline{\gamma}^h$ and $V^h(\cdot)$ are used to denote the solution to the discretized equation and we use the definitions $g^+(x) = \max[g(x),0]$ and $g^-(x) = \max[0,-g(x)]$.

(13) $$h^2\bar{\gamma}^h = \inf_{\alpha \varepsilon \mathcal{U}} [-Q_h(x,\alpha)V^h(x) + \sum_{i,\pm} V^h(x \pm e_i h)(hf_i^{\pm}(x,\alpha) + \sigma_i^2(x)/2) + h^2 k(x,\alpha)], \quad x \varepsilon G_h,$$

$$h\bar{\gamma}^h = \inf_{\alpha \varepsilon \mathcal{U}_0} [-Q_h(x,\alpha)V^h(x) + \sum_{i,\pm} V^h(x \pm e_i h)\gamma_i^{\pm}(x,\alpha) + hk_0(x,\alpha)], \quad x \varepsilon \partial G_h.$$

Define $p^h(x,x \pm e_i h|\alpha) = (\text{coefficient of } V^h(x \pm e_i h))/\bar{Q}_h(x)$, $p^h(x,x|\alpha) = [\bar{Q}_h(x) - Q_h(x,\alpha)]/\bar{Q}_h(x)$. Divide (13) through by $\bar{Q}_h(x)$ and rearrange to get

(14) $$V^h(x) + \bar{\gamma}^h \Delta t^h(x) = \inf_{\alpha \varepsilon \mathcal{U}} [\sum_{i,\pm} V^h(x \pm e_i h)p^h(x,x \pm e_i h|\alpha) + V^h(x)p^h(x,x|\alpha) + k(x,\alpha)\Delta t^h(x)], \quad x \varepsilon G_h,$$

and similarly for x in ∂G_h, where $\mathcal{U}$ and k are replaced by $\mathcal{U}_0$ and k_0, resp. Define $p^h(x,y|\alpha) = 0$ for all x,y other than $y = x$ or $y = x \pm e_i h$ for some i. Then $\{p^h(x,y|\alpha),\ x,\ y \varepsilon \bar{G}_h\}$ is a transition probability for a controlled Markov chain. Let $\{\xi_n^h\}$ denote the random variables of the chain, and define $\mathcal{U}(x) = \mathcal{U}$ in G, and $\mathcal{U}(x) = \mathcal{U}_0$ on ∂G, and <u>redefine</u> $k(x,\alpha)$ to equal $k_0(x,\alpha)$ for $x \varepsilon \partial G$. Then (14) can be rewritten in the form

(15) $$V^h(x) + \bar{\gamma}^h \Delta t^h(x) = \inf_{\alpha \varepsilon \mathcal{U}(x)} [E_x^{\alpha} V^h(\xi_1^h) + k(x,\alpha)\Delta t^h(x)], \quad x \varepsilon \bar{G}_h.$$

In (13)-(15), we supposed that $\bar{\gamma}^h$ is a constant. This is almost equivalent to the assumption that there is only one recurrence class for the chain under the optimal control. If there is more than one recurrence class, the numerical problem is harder. Let us henceforth assume

(A2) <u>For each small h and under each stationary pure Markov control, there is only one recurrence class.</u>

This assumption seems to hold in very many cases of practical interest. It can be dispensed with, but then the problem of actually solving (13)-(15) is much harder. Under (A2), (15) can be solved by either Howard's iteration in policy space for semi-Markov processes, or by a version of the backward iteration method for the average cost per unit time problem (see, e.g., Schweitzer and Federgruen [8], but adapted to a semi-Markov process model). There is an optimal stationary pure Markov control $u^h(\cdot)$ for all small h, it is the minimizer in (15), and it is optimal with respect to all controls for the discrete problem. The "Semi-Markov" point will be returned to below. The optimal solution is given in the first line of (19).

Discussion of (14). For $y \in G_h$, we have for any stationary pure Markov control $u(\cdot)$

$$\text{(16a)} \qquad E_y^u[\xi_{n+1}^h - \xi_n^h \mid \xi_n^h = y,\ u(\cdot) \ \text{used}] = f(y,u(y))\Delta t^h(y),$$

$$\text{cov}_y^u[\xi_{n+1}^h - \xi_n^h \mid \xi_n^h = y,\ u(\cdot) \ \text{used}] = \sigma(y)\sigma'(y)\Delta t^h(y) + o(\Delta t^h(y)), \quad x \in G_h.$$

For $y \in \partial G_h$,

$$\text{(16b)} \qquad \begin{aligned} E_y^u[\xi_{n+1}^h - \xi_n^h \mid \xi_n^h = y,\ u(\cdot) \ \text{used}] &= \gamma(y,u(y))\Delta t^h(y), \\ \text{cov}_y^u[\xi_{n+1}^h - \xi_n^h \mid \xi_n^h = y,\ u(\cdot) \ \text{used}] &= o(\Delta t^h(y)). \end{aligned}$$

These "infinitesimal" properties (derived in [2], [3]), together with (15), suggest a close relation between the controlled chain, and the controlled reflected diffusion.

These relations are brought out quite clearly when the chain is suitably interpolated into a continuous parameter process, and (15), (16) suggest several useful interpolations. First, we note that solving (15) is the only computation that need be done. Equation (15) is not quite the dynamic programming equation for the average cost per unit time for the controlled chain $\{\xi_n^h\}$, since $\bar{\gamma}^h$ has a state dependent coefficient $\Delta t^h(\cdot)$. However, it is the dynamic programming equation for a semi-Markov process or, equivalently for the types of continuous parameter interpolations which are discussed below.

Let π^h denotesthe invariant measure which corresponds to the optimal control. Henceforth, unless otherwise mentioned, $\{\xi_n^h\}$ refers to the optimal chain, with initial measure π^h.

We now choose an interpolation method and show that the sequence of interpolated processes converges weakly to a solution to the submartingale problem corresponding to some admissible control

$u(\cdot)$, and that this solution is an optimal one, with cost rate $\bar{\gamma} = \lim_{h\to 0} \bar{\gamma}^h$.

Either of the following two piecewise constant interpolations will serve our purpose.

<u>Interpolation 1</u>. Define $\Delta t^h(\xi_i^h) = \Delta t_i^h$, $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$. Define the semi-Markov process $\xi^h(\cdot)$ by $\xi^h(t) = \xi_n^h$ on $[t_n^h, t_{n+1}^h)$. This interpolation was used in [2], [3].

<u>Interpolation 2</u>. Let $\xi^h(\cdot)$ denote the Markov jump process on $\bar{G}_h$ defined by:

If $\xi^h(t) = y$, then the average additional time spent in state y before a jump is $\Delta t^h(y)$, and $P\{\text{next state} = y' \mid \text{current state} = y\} = p^h(y,y'|u^h(y))$. There is a slight ambiguity here since it is possible that $p^h(y,y|u^h(y)) > 0$. But, this should cause no confusion - for it simply means that there is a jump of "zero" magnitude. The average interjump times can be normalized to avoid this, but it hardly seems worthwhile. Note that

$$P\{\text{jump in } (t,t+\Delta] | \xi^h(t) = y\} = (\Delta/\Delta t^h(y)) + o(\Delta).$$

This interpolation is developed in Section 8 of [4].

Neither interpolation is always preferable to the other. Interpolation 2 could have been used in references [2], [3], but there did not seem to be a need for it then. There are advantages to having an interpolation which is a continuous parameter Markov chain in that certain concepts (such as stationarity) have a clearer meaning; on the other hand it is sometimes preferable to work with interpolation times that are deterministic functions of the current state, since then there are fewer random variables to worry about. The limiting processes (see Sections 6 and 7) are the same for both interpolations. In Case 2, the average sojourn time in a state y (before the next jump, whether of zero value or not) is $\Delta t^h(y)$, precisely the interpolation interval for Case 1. In both cases, the time spent at a state y on the boundary ($O(h)$, per sojourn) is greater than time spent at a state y in G_h ($O(h^2)$ per sojourn, unless there is the complete degeneracy $\sigma(y) = 0$). This property is a consequence of our definition of $\Delta t^h(y)$ for $y \in \partial G_h$

(to correspond to $\rho(y) \equiv 1$).

For either Interpolation 1 or 2,

$$\bar{\gamma}^h = \lim_{t\to\infty} E_x^h \int_0^t k(\xi_s^h, u_s^h)\,ds/t, \tag{17}$$

where $u_s^h = u^h(\xi_s^h)$, and E_x^h indicates that u^h is used. The invariant measure for either interpolation is μ^h, where

$$\mu^h(y) = \Delta t^h(y)\pi^h(y)/\sum_z \Delta t^h(z)\pi^h(z) \tag{18a}$$

Also,

$$\bar{\gamma}^h = \sum_y \mu^h(y)k(y,u^h(y)). \tag{18b}$$

Equations (17) and (18) are not hard to verify. For example, (18) follows from the ergodic theorems for Markov chains (see Chung [6], Section 1.15, Theorems 1, 2, 3; see also [2], Chapter 6.8, for similar calculations). It can also be obtained by direct verification of the Kolmogorov equation using the invariance of $\pi^h(\cdot)$ for the discrete parameter chain. To get (17) write u_i^h for $u^h(\xi_i^h)$ and use (15) and the same ergodic theorems to get

$$\begin{aligned}
\bar{\gamma}^h &= \lim_{n\to\infty} [E_x^h \sum_{i=0}^{n-1} k(\xi_i^h, u_i^h)\Delta t_i^h / E_x^h \sum_{i=0}^{n-1} \Delta t_i^h] \\
&= \lim_{\substack{n \\ (\text{w.p.1})}} [\sum_{i=0}^{n-1} k(\xi_i^h, u_i^h)\Delta t_i^h / \sum_{i=0}^{n-1} \Delta t_i^h] \\
&= \lim_{\substack{n \\ (\text{w.p.1})}} \int_0^{t_n^h} k(\xi_s^h, u_s^h)\,ds/t_n^h = \lim_{\substack{(\text{w.p.1}) \\ t\to\infty}} \int_0^t k(\xi_s^h, u_s^h)\,ds/t \\
&= \lim_{t\to\infty} \int_0^t E_x^h k(\xi_s^h, u_s^h)\,ds/t.
\end{aligned} \tag{19}$$

Similarly, the first limit in (19) equals

$$\begin{aligned}
\bar{\gamma}^h &= \sum_y \pi^h(y)k(y,u^h(y))\Delta t^h(y)/\sum_y \pi^h(y)\Delta t^h(y) \\
&= \sum_y \mu^h(y)k(y,u^h(y)).
\end{aligned} \tag{20}$$

Let $v(\cdot)$ denote a stationary pure Markov control. Then (15) implies that (here $\Delta t_i^h, \xi_i^h$ now refer to the variables under control $v(\cdot)$) for any x

$$\text{(21)}\qquad \bar{\gamma}^h \le \lim_{n\to\infty} \frac{E_x^v \sum_{i=0}^{n-1} \Delta t_i^h k(\xi_i^h, v(\xi_i^h))}{E_x^v \sum_{i=0}^{n-1} \Delta t_i^h} \equiv \gamma^{v,h}.$$

The proof of optimality of $u^h(\cdot)$ with respect to any control which is not necessarily stationary pure Markov can be based on a method of Ross [7] and is omitted.

6. Weak Convergence. We will work with Interpolation 2, since it is a strictly stationary process. The method will be outlined, but the proofs will be usually referred to when already available elsewhere. So far, we have a sequence of stationary pure Markov controls $\{u^h(\cdot)\}$, corresponding stationary continuous parameter Markov chains $\{\xi^h(\cdot)\}$, invariant measures $\{\mu^h\}$, and minimum costs $\{\bar{\gamma}^h\}$, where

$$\bar{\gamma}^h = \sum_{y\in\bar{G}_h} \mu^h(y)k(y,u^h(y)) = \sum_{y\in G_h} \mu^h(y)k(y,u^h(y))$$

$$+ \sum_{y\in\partial G_h} \mu^h(y)k_0(y,u^h(y)),$$

and

$$\text{(22)}\qquad \bar{\gamma}^h t = E^h\left[\int_0^t k(\xi_s^h,u_s^h)I_G(\xi_s^h)ds + \int_0^t k_0(\xi_s^h,u_s^h)I_{\partial G}(\xi_s^h)ds\right],$$

where E^h denotes the expectation under initial measure μ^h, and we use $u_s^h = u^h(\xi_s^h)$. We often write $\xi^h(s)$ as ξ_s^h, etc., for typographical simplicity.

We obviously can write

$$\text{(23)}\qquad \xi_s^h = \xi_0^h + \int_0^t I_G(\xi_s^h) f(\xi_s^h,u_s^h)ds$$

$$+ \int_0^t I_{\partial G}(\xi_s^h)\gamma(\xi_s^h,u_s^h)ds + B^h(t) + B_0^h(t),$$

where

$$B^h(t) = \int_0^t I_G(\xi_s^h)[d\xi^h(s) - f(\xi_s^h,u_s^h)ds],$$

$$B_0^h(t) = \int_0^t I_{\partial G}(\xi_s^h)[d\xi_s^h - \gamma(\xi_s^h,u_s^h)ds].$$

Denote the two integrals in (22) by $K^h(t)$ and $K_0^h(t)$, resp., and the first two integrals on the right side of (23) by $Q^h(t)$ and $Q_0^h(t)$, resp. Let $D^m[0,\infty)$ denote the space of R^m valued functions on $[0,\infty)$, continuous on the right and with left-hand limits (Billingsley [9], Lindvall [10], Kushner [2], Chapter 2), endowed with the Skorokhod topology. If a measure ν_n induces a process $X^n(\cdot)$ with paths in $D^m[0,\infty)$ w.p.1 and $\{\nu_n\}$ is tight, we abuse terminology and say that $\{X^n(\cdot)\}$ is tight. If $\{\nu_n\}$ converges weakly to a measure ν and ν induces a process $X(\cdot)$ with paths in $D^m[0,\infty)$ w.p.1, we say that $\{X^n(\cdot)\}$ converges weakly to $X(\cdot)$. We occasionally use Skorokhod imbedding ([11], Theorem 3.1.1, or [2], Chapter 2), which says that if $X^n(\cdot) \to X(\cdot)$ weakly in $D^m[0,\infty)$, then there are processes $\tilde{X}(\cdot),\tilde{X}^n(\cdot)$ with paths in $D^m[0,\infty)$ and which induce the same measures on $D^m[0,\infty)$ as do $X(\cdot),X^n(\cdot)$, resp., and are such that $\tilde{X}^n(\cdot) \to \tilde{X}(\cdot)$ w.p.1 in the Skorokhod topology. Since all our limit processes will be continuous w.p.1, this implies that $\tilde{X}^n(t) \to \tilde{X}(t)$, uniformly on bounded intervals. Also, we omit the tilde ~ notation. The following theorem follows from the results in [4], Section 8.

<u>Theorem 1.</u>[+] $\{\xi^h(\cdot), K^h(\cdot), K_0^h(\cdot), B^h(\cdot), B_0^h(\cdot), Q^h(\cdot), Q_0^h(\cdot)\} \equiv \{\Phi(\cdot)\}\}$ <u>is tight on</u> $D^{5r+2}[0,\infty)$, <u>and all limits have continuous paths</u> w.p.1.

We will next characterize the limits of $\{B^h(\cdot),B_0^h(\cdot)\}$.

Let us choose a weakly convergent subsequence, also indexed by h, and henceforth fixed. The subsequent results will not depend upon the selected subsequence. Denote the limit by $\xi(\cdot)$, $K(\cdot)$, $K_0(\cdot)$, $B(\cdot)$, $B_0(\cdot)$, $Q(\cdot)$, $Q_0(\cdot)$. By construction, $B^h(t)$ and

---

[+]Theorem 1 does not require A1 or A2 and holds whether the initial conditions are random or not. It needs only the boundedness and continuity of f,σ,k,k_0 and γ. Also, u^h can be replaced by any pure Markov control.

$B_0^h(\cdot)$ are martingales (with respect to the σ-algebras B_t^h induced by ξ_s^h, $s \leq t$) and an easy calculation yields that

$$E \sup_{t \leq T} |B_0^h(t)|^2 \leq \text{constant} \cdot hT.$$

Thus $B_0(\cdot)$ is the zero process.

The quadratic variation of $B^h(\cdot)$ is

$$\int_0^t \Sigma^h(\xi_s^h) I_G(\xi_s^h) ds,$$

where $\Sigma^h(x)$ is such that it converges to $\sigma(x)\sigma'(x)$ as $h \to 0$, uniformly in x, and $\sup_h E|B^h(t)|^4 < \infty$ for each $t > 0$. Then $\{|B^h(t)|^2\}\}$ is uniformly integrable for each t. Let $\mathscr{B}_t$ denote the σ-algebra induced by $\{\xi_s, B(s), K(s), K_0(s), Q(s), Q(s),\ s \leq t\}$. Let N_ε denote an ε neighborhood of ∂G. In [3], Lemma 1, it is shown that for each real $T > 0$ there is a constant K such that, for Interpolation 1 and small $\varepsilon > 0$

$$E_x^u \int_0^T I_{N_\varepsilon}(\xi_s^h) I_G(\xi_s^h) ds \leq K_T \varepsilon, \tag{24}$$

uniformly in u,h (although u did not appear in the derivation, only an upper bound to the values of the drift function f was used in the derivation). The result (24) depends only on the fact that the component of the diffusion term $\sigma(x)dw$ orthogonal to the boundary is uniformly non-degenerate on ∂G; i.e. on (A1). Estimate (24) also holds for Interpolation 2, and is crucial for the rest of the development. It says that neither the approximations nor the limit can "linger" near (but not on) the boundary. In particular, it implies that the probability is zero that over some subinterval of $[0,T]$ the paths for the approximations will be in $N_\varepsilon \cap G$ and the limit will be on ∂G.

<u>Theorem 2</u>. <u>Assume A1</u>. $\{B(t), \mathscr{B}_t\}$ <u>is a continuous martingale with quadratic covariation</u> $\int_0^t I_G(\xi_s)\sigma(\xi_s)\sigma'(\xi_s)ds$.

<u>Proof</u>. The proof, using (24), follows similar calculations in [2], [3], [4]. Let $q^h(t)$ represent any of the vectors in $\phi^h(\cdot)$ (see Theorem 1), let n denote an arbitrary integer, t_i, $i \leq n$, numbers less than or equal to t, let $s > 0$ and let $g(\cdot)$ denote

a continuous real valued function. By weak convergence, Skorokhod imbedding and the uniform integrability of $\{|B^h(t)|\}$ for each t, the result (martingale property of $B^h(\cdot)$)

$$E^h g(q^h(t_i), i \le n)[B^h(t+s) - B^h(t)] = 0$$

implies

$$Eg(q(t_i), i \le n)[B(t+s) - B(t)] = 0.$$

Also, the result

$$E^h g(q^h(t_i), i \le n)[(B^h(t+s) - B^h(t))(B^h(t+s) - B^h(t))' - \int_0^t I_G(\xi_s^h)\Sigma^h(\xi_s^h)ds] = 0$$

together with the weak convergence, Skorokhod imbedding and uniform+ integrability of $\{|B^h(t)|^2\}$ and (24) implies that

$$Eg(q(t_i), i \le n)[(B(t+s) - B(t))(B(t+s) - B(t))' - \int_0^t I_G(\xi_s)\sigma(\xi_s)\sigma'(\xi_s))ds] = 0.$$

The arbitrariness of $g(\cdot)$, t, $t + s$, t_i, $i \le n$, and n imply the theorem. Q.E.D.

We next need a representation for $Q(\cdot)$, $Q_0(\cdot)$, $K(\cdot)$ and $K_0(\cdot)$. It is easy to see that all these functions are absolutely continuous with respect to Lebesgue measure. Thus, there are measurable (ω,t) functions $\bar{q}(\cdot)$, $\bar{q}_0(\cdot)$, $\bar{k}(\cdot)$ and $\bar{k}_0(\cdot)$ such that, for almost all ω,t,

$$Q(t) = \int_0^t \bar{q}(s)ds, \qquad Q_0(t) = \int_0^t \bar{q}_0(s)ds$$

$$K(t) = \int_0^t \bar{k}(s)ds, \qquad K_0(t) = \int_0^t \bar{k}_0(s)ds.$$

+Actually, uniform integrability of $\{|B^h(t)|^2\}$ (implied by $\sup_h E^h|B^h(t)|^4 < \infty$ is not needed. Since $B(\cdot)$ is a square integrable continuous martingale, its quadratic variation can be obtained by a "localization" of the argument.

We can now proceed in two ways, either working with generalized random controls or by imposing a convexity condition and using an implicit function theorem. We take the latter (and easier) approach.

Theorem 3.[+] Assume A1 and A2. Let $f, k, k_0, \gamma, \sigma$ be continuous and let the sets $\{f(x,\alpha), k(x,\alpha), \alpha \in \mathscr{U}\} \equiv g(x,\mathscr{U})$ and $\{\gamma(x,\alpha), k_0(x,\alpha), \alpha \in \mathscr{U}_0\} \equiv g_0(x,\mathscr{U}_0)$ be convex for each x. Then there is a control $\bar{u}(\cdot)$[+] with values $\bar{u}_s$ in $\mathscr{U}$ when $\xi_s \in G$ and in $\mathscr{U}_0$ when $\xi_s \in \partial G$ and such that, for almost all ω, t,

$$\bar{f}(t) = f(\xi_t,\bar{u}_t) I_G(\xi_t)$$

$$\bar{f}_0(t) = \gamma(\xi_t,\bar{u}_t) I_{\partial G}(\xi_t)$$

$$\bar{k}(t) = k(\xi_t,\bar{u}_t) I_G(\xi_t)$$

$$\bar{k}_0(t) = k_0(\xi_t,\bar{u}_t) I_{\partial G}(\xi_t).$$

Proof. Define $\bar{g}(t) = (\bar{f}(t),\bar{k}(t))$ and $\bar{g}_0(t) = (\bar{f}_0(t),\bar{k}_0(t))$. The proof uses the basic estimate (24) and the method of [2], pp. 182-183. By (24) and [2], pp. 182-183, for almost all ω, t

$$\bar{g}(t) \in g(\xi_t, \mathscr{U}) I_G(\xi_t)$$

$$\bar{g}_0(t) \in g_0(\xi_t, \mathscr{U}_0) I_{\partial G}(\xi_t),$$

from which the result follows by the McShane-Warfield implicit function theorem as in [2], Theorem 9.2.2. Q.E.D.

Summing up the results of Theorems 1 to 3, we get the representation (under A1 and A2)

$$(25) \quad \xi_t = \xi_0 + \int_0^t I_G(\xi_s) f(\xi_s,\bar{u}_s)ds + \int_0^t I_{\partial G}(\xi_s)\gamma(\xi_s,\bar{u}_s)ds + B(t),$$

where $B(t)$ is a continuous martingale with quadratic variation

$$\int_0^t I_G(\xi_s)\sigma(\xi_s)\sigma'(\xi_s)ds.$$

[+]This control is also non-anticipative with respect to the $w(\cdot)$ introduced below (25).

Also, there is a Wiener process $w(\cdot)$, with respect to which all the other processes in (25) are non-anticipative and such that $B(t) = \int_0^t I_G(\xi_s)\sigma(\xi_s))dw(s)$. Obviously, by the weak convergence, ξ_t is in $\overline{G}$ for all t. Let $\mathcal{L}^{\overline{u}}$ denote the differential generator associated with (25) in G. By a slight modification of the argument associated with (40) and (41) in [3], we can show that $\xi(\cdot)$ solves the sub-martingale problem.

Furthermore, $\xi(\cdot)$ is a stationary process. Let its invariant measure be denoted by μ, (which is the weak limit of $\{\mu^h\}$), and let $\overline{\gamma} = \lim_h \overline{\gamma}^h$. Then the distribution of ξ_0 is μ. By (22), (24),

$$\overline{\gamma}t = E^{\overline{u}}[\int_0^t I_G(\xi_s)k(\xi_s,\overline{u}_s)ds + \int_0^t I_{\partial G}(\xi_s)k_0(\xi_s,\overline{u}_s)ds]. \tag{26}$$

<u>Remarks</u>. The limit process $\xi(\cdot)$ is stationary, as is the drift $\overline{f}(\cdot)$, but we have not been able to show that there is a Markov (reflecting diffusion) process with the same distributions. There probably is such a Markov process, as there probably is a stationary pure Markov control $\tilde{u}(\cdot)$ such that $\tilde{u}(\xi_t) = \overline{u}(\omega,t)$ w.p.1. In any case, our method gives much information on the optimal process $\xi(\cdot)$; e.g., the multivariate distribituions of $\xi^h(\cdot)$ converge weakly to those of $\xi(\cdot)$, as do the distributions of any bounded measurable functional $F(\xi^h(\cdot))$, if $F(x(\cdot))$ is continuous w.p.1 with the respect to the measure induced by $\xi(\cdot)$. Indeed, one of the great advantages of the weak convergence method is that it yields such information, in addition to approximations to $\overline{\gamma}$. Also, $\overline{\gamma}$ = average cost per unit time for $\xi(\cdot)$, and is the limit of the average costs per unit time for the sequence of approximations.

7. <u>Optimality of the Limit</u> $\xi(\cdot)$. Being a limit of optimal approximating processes, $\xi(\cdot)$ is a good candidate for optimality for the original optimization problem (with the reflected diffusion model). Certain optimality properties are easy to show.

<u>Theorem 4</u>. <u>Assume A1 and A2</u>. <u>Let</u> $v(\cdot)$ <u>denote a continuous stationary pure Markov control, such that the corresponding reflecting diffusion</u> $\xi^v(\cdot)$ <u>is unique (in the weak sense) and has a</u>

<u>unique invariant measure</u> μ^v. <u>Then</u> $\bar{\gamma} \le \gamma^v$ <u>(where we let the initial measure be</u> μ^v).

<u>Proof</u>. Let $\hat{\xi}_n^h$ and $\hat{\xi}^h(\cdot)$ denote the discretized and interpolated processes, resp., corresponding to the fixed control $v(\cdot)$. Then the cost $\gamma^{v,h}$ for the interpolated process is $\ge \bar{\gamma}^h$ by optimality of u^h. Let $\mu^{v,h}$ denote <u>any</u> invariant measure for $\hat{\xi}^h(\cdot)$. Then $\{\hat{\xi}^h(\cdot)\}$ and the invariant measures $\{\mu^{v,h}\}$ converge weakly to $\xi^v(\cdot)$ and μ^v, resp., as $h \to 0$ by arguments similar to those in Theorems 1 to 2. The theorem follows from this and (24).

Q.E.D.

Since we have not been able so far to prove that $\bar{u}(\cdot)$ is stationary pure Markov, it would be nice to prove that $\bar{u}(\cdot)$ is optimal with respect to a broader class of controls than those in Theorem 4. The class can be broadened, but at the expense of considerable terminology and detail. We refer the reader to [2], where broader classes of comparison controls are dealt with for a number of other types of optimization problems.

REFERENCES

[1] Kushner, H.J. (1978) Optimality conditions for the average cost per unit time problem with a diffusion model, SIAM J. Control and Optimization, 16, pp. 330-346.

[2] Kushner, H.J. (1977) Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York.

[3] Kushner, H.J. (1976) Probabilistic methods for finite difference approximations to degenerate elliptic and parabolic equations with Neumann and Dirichlet boundary conditions, J. Math. Anal. and Applic., 53, pp. 643-688.

[4] Kushner, H.J. and DiMasi, G (1978) Approximations for functionals and optimal control problems on jump diffusion processes, J. Math. Anal. and Applic., 63, pp. 772-800.

[5] Strook, D.W. and Varadhan, S.R.S. (1971) Diffusion processes with boundary conditions, Comm. Pure Appl. Math., 24, pp. 147-225.

[6] Chung, K.L. (1960) Markov Chains with Stationary Transition Probabilities, Springer-Verlag, Berlin.

[7] Ross, S.M. (1968) Arbitrary state Markovian decision processes, Ann. Math. Stat. 39, pp. 2118-2122.

[8] Schweitzer, P.J. and Federgruen, A. (1977) The asymptotic behavior of undiscounted value iteration in Markov decision problems, Math. of Oper. Res. 2, pp. 360-381.

[9] Billingsley, B. (1968) Convergence of Probability Measures, Wiley, New York.